CONTRIBUTIONS FROM THE ZOÖLOGICAL LABORATORY OF THE MUSEUM OF COMPARATIVE ZOÖLOGY AT HARVARD COLLEGE, E. L. MARK, DIRECTOR, NO. LXXIX.

A MEASURE OF VARIABILITY, AND THE RELATION OF INDIVIDUAL VARIATIONS TO SPECIFIC DIFFERENCES.

BY EDWIN TENNEY BREWSTER.

Presented by E. L. Mark, April 14, 1897.

PROBLEM.

THIS paper, prepared under the supervision of Dr. C. B. Davenport, deals with an inquiry into the relation between those small variations which distinguish individuals of the same group, and those larger differences which separate species and genera.

For the prosecution of such an inquiry, it is, first of all, necessary to devise a method of measuring variability. Quetelet ('46), Stieda ('82), Galton ('91), and Weldon ('93) have shown that variations in organisms follow the well known laws of the distribution of error. Thus the ordinary methods of treating problems in error of observation may be made to furnish a measure of variability; it is upon these methods that the methods of this paper are based.

METHOD.

Any measurable quality of an object has a value, which is expressible by a number. A series of such numbers, expressing the varying value of a quality throughout a group of similar objects, will have a mean, about which the quantities are arranged in accordance with the law of distribution of error. Such a series is, for the present purpose, essentially like a series of slightly erroneous observations of a single quantity distributed about the true value. What, therefore, would be the probable error in the latter case, is a measure of variability in the former.

Suppose such a series of numbers obtained by measuring some single quality in each individual of a natural group of organisms. Let there be

n individuals, and let $d_1, d_2 \ldots d_n$ represent the difference between each number of the series and the mean of all.

Then

$$\frac{0.8453\ (d_1 + d_2 \ldots d_n)}{\sqrt{n\ (n-1)}},$$

which is the common working formula for probable error,* is a measure of the variability of the given quality in this particular group.

Or, in symbols,

$$V \propto \frac{0.8453\ \Sigma\, d}{\sqrt{n\ (n-1)}};$$

where V stands for variability, $\Sigma\, d = d_1 + d_2 \ldots d_n$, and $\propto$ may be read "is measured by." This is, appproximately, Galton's Q.

Obviously,

$$V \propto \frac{\Sigma\, d}{\sqrt{n\ (n-1)}}. \qquad \text{(Formula 1.)}$$

In Formula 1,

$$\sqrt{n\ (n-1)} = \sqrt{n^2 - n} = \sqrt{n^2\left(1 - \frac{1}{n}\right)} = n\sqrt{1 - \frac{1}{n}}.$$

$$\therefore\ V \propto \frac{\Sigma\, d}{n} \times \frac{1}{\sqrt{1 - \frac{1}{n}}}.$$

Consider the expression, $\frac{\Sigma\, d}{n}$.

Here $\Sigma\, d$ is the sum of all differences between single numbers of the series and the mean of the series. Since these are n in number, $\frac{\Sigma\, d}{n}$ is the *average deviation* of single values from the mean value.

Let $a\, d$ be the symbol for this average deviation. Then,

$$V \propto a\, d \frac{1}{\sqrt{1 - \frac{1}{n}}}. \qquad \text{(Formula 2.)}$$

Suppose, however, that the number of cases measured is large; that is to say, that n is made indefinitely large. As n approaches infinity, $\frac{1}{n}$ approaches 0, and consequently $\frac{1}{\sqrt{1 - \frac{1}{n}}}$ approaches 1.

---

* See any text-book treating of such subjects; for example, Merriman ('84), p. 93.

If, then, n be taken sufficiently large,

$$V \propto a\,d. \qquad \text{(Formula 3.)}$$

Even if n is not large, Formula 3 may usually be employed in practice. For in any single investigation n is likely to be constant, — when, as will often be the case, different qualities of the same individuals are to be compared it is necessarily so, — but the introduction of a constant factor will not affect the correctness of the formula.

Practically, $a\,d$ may be found accurately enough without the labor of subtracting each quantity from the mean.

Let

$$a + \frac{n}{2}, \dots\, a + d,\ a + c,\ a + b,\ a - b,\ a - c,\ a - d \dots\, a - \frac{n}{2}$$

be a series of quantities, n in number, distributed according to the law of probability. The mean is evidently a, and the differences between the several quantities and the mean are

$$\frac{n}{2}, \dots\, d,\ c,\ b,\ b,\ c,\ d, \dots \frac{n}{2}.$$

$$\therefore\ a\,d = \frac{\frac{n}{2} \dots + d + c + b + b + c + d \dots + \frac{n}{2}}{n}$$

$$= \frac{2\,b + 2\,c + 2\,d \dots + n}{n}. \qquad \text{(Formula 4.)}$$

Or, again,

$$a + \frac{n}{2}, \dots\, a + d,\ a + c,\ a + b,\ a - b,\ a - c,\ a - d, \dots\, a - \frac{n}{2},$$

being the series of n measurements as before, the sum of all terms greater than the mean is

$$\left(a + \frac{n}{2}\right) \dots + (a + d) + (a + c) + (a + b),$$

and of these the mean is

$$\frac{\frac{n\,a}{2} + \frac{n}{2} \dots + d + c + b}{\frac{n}{2}}.$$

In like manner, the sum of all terms less than the mean is

$$\frac{\frac{n\,a}{2} - \frac{n}{2} \ldots\ldots - d - c - b}{\frac{n}{2}}.$$

One half the difference between these is

$$\frac{2\,b + 2\,c + 2\,d \ldots\ldots n}{n}. \qquad \text{(Formula 5.)}$$

But Formula 5 is identical with Formula 4, and is therefore the formula for the average deviation.

Hence the average deviation may be found by separating the numbers into two groups, — one of which shall contain all quantities greater than the mean, the other all those less than the mean, — and taking half the difference between the means of each group.

In practice there will usually be several observed values equal to the mean; these may be distributed to make the two groups of equal size. If the number of observed values is odd, one of these mean values may be neglected.

In order to compare the average deviation of one group of numbers with that of another which has a different mean, it is necessary to reduce the two to a common measure. This is most simply done by dividing each average deviation by the corresponding mean.* It is proposed to call this ratio the *coefficient of variability* and to designate it by the symbol *C. V.*

---

* The justification for this procedure is found in the following considerations: The relative size of the average deviation of two organs depends very largely upon the relative size of these organs. Where the mean dimension is large, we expect a greater average deviation than where it is small. Thus the average deviation of the stature of adult British males from the mean is about 2 inches. An average deviation of 2 inches in the length of the nose, in any race, would clearly indicate a much greater variability in the nose length than in stature. In comparing the variability of two such diverse measures as stature and nose length, it is better to compare the ratios of the average deviations to the mean dimension. Thus, since the mean stature of adult British males may be taken at 67 inches, variability in stature may be expressed by the ratio $\frac{2}{67} = .02985$. This number indicates that the average deviation from the mean stature is about three one-hundredths of the mean stature; which is clearly more important than to say that it is 2 inches. Moreover, this method of expression has the advantage that it is independent of the unit in which the dimension is measured, whether feet, millimeters, grams, degrees, or ergs. — C. B. DAVENPORT.

In practice *C. V.* may be found most readily by separating the given numbers into two groups, as discussed above, and finding the mean of each. The difference between these two means, divided by their sum, is the *C. V.* of the structure under consideration.

Application of the Method of Measuring Variability to the Problem of the Relation between Individual Variations and Specific Differences.

Thesis.

While it is generally agreed that *specific* characters are more subject to striking variations in individuals than are the characters *common to allied species,* it is not clear how far this relation extends. I wish to show that what is true of obvious variations and sports is also true of those minute differences between individuals which only careful measurements can detect; or, in other words, that any measurable quality is, in general, variable in individuals in proportion as it is a distinguishing character of the group to which the individuals belong.

Evidence.

Table A is based on the extensive tables of body measurement of twenty races of men, which are given by Weisbach ('78) in the Appendix to his "Körpermessungen," etc. Each of the numbers of the first eight columns is the *C. V.* of a single dimension of a single race, and the ninth column gives the average for eight races of the *C. V.*'s of each dimension. The values of these *C. V.*'s only approximate to the true values, for the reason that the number of individuals measured is not sufficiently large to eliminate all accidents of age and sex.

An examination of Table A shows that these eighteen dimensions are of nearly the same relative variability in each of the eight races. This fact is well brought out by Table B, in which the largest number of each column of Table A is replaced by 1, the next smaller by 2, and so on up to 18. From Table B it appears that certain dimensions — as, for example, the height of the forehead — are always decidedly variable; and, on the other hand, it appears that other dimensions — such as the length and breadth of the head — are more constant.

The last column of Table A, "Mean of 20 Races," gives the coefficient of variability, not of any individuals, but of the means of each of the twenty races of Weisbach's tables. That is, the *mean value* of a dimension in each race is treated as are the dimensions of individuals in other columns. This column, therefore, shows the distribution of *racial* differences in the same way in which the remainder of the table shows

TABLE A.

C. V. OF CERTAIN DIMENSIONS IN VARIOUS HUMAN RACES.

(Based on Weisbach, '78, Appendix.)

	24 Slavs.	8 Siamese.	12 Japanese.	20 Chinese.	9 Kanakas.	19 Jews.	20 Magyars.	26 Roumanians.	Average.	Mean dimensions of each of 20 Races.
Head length	0.0314	0.021	0.0205	0.028	0.030	0.0309	0.0293	0.0302	0.0277	0.0244
Head breadth	0.0345	0.024	0.0289	0.287	0.021	0.0363	0.0322	0.0338	0.0298	0.0278
Nose length	0.0532	0.051	0.0504	0.0836	0.041	0.0435	0.17	0.0633	0.0695	0.0949
Nose breadth	0.0446	0.095	0.051	0.064	0.096	0.0447	0.0546	0.0481	0.0613	0.0757
Nose height	0.0683	0.116	0.087	0.106	0.049	0.0575	0.0625	0.0601	0.0758	0.152
Forehead height	0.0721	0.078	0.105	0.087	0.087	0.0941	0.0819	0.095	0.0778	0.104
Under jaw length	0.0459	0.0296	0.0427	0.0567	0.0501	0.0539	0.0435	0.0322	0.0443	0.0481
Mouth breadth	0.0619	0.089	0.0856	0.048	0.0374	0.0619	0.0550	0.0496	0.0611	0.0518
Upper face breadth	0.0457	0.0565	0.0512	0.030	0.0432	0.0361	0.0392	0.0354	0.0422	0.0322
Lower face breadth	0.0288	0.0335	0.047	0.038	0.0678	0.0571	0.0528	0.0344	0.0449	0.0494
Upper arm length	0.0475	0.045	0.043	0.0515	0.0439	0.0423	0.0396	0.0310	0.0430	0.0650
Forearm length	0.0447	0.0259	0.0282	0.0504	0.026	0.0340	0.0412	0.0353	0.0357	0.0385
Middle finger length	0.0537	0.0876	0.0391	0.0429	0.0526	0.0501	0.0297	0.0344	0.0488	0.0595
Hand width	0.0450	0.0344	0.0299	0.0506	0.0493	0.0395	0.0334	0.0236	0.0382	0.0793
Upper leg length	0.0324	0.0391	0.0395	0.0606	0.0347	0.0318	0.0416	0.0310	0.0388	0.0500
Lower leg length	0.0412	0.0205	0.045	0.0591	0.0528	0.0375	0.0350	0.0312	0.0403	0.0504
Foot length	0.0441	0.0620	0.031	0.0238	0.0428	0.0351	0.0267	0.0280	0.0367	0.0592
Foot breadth	0.0499	0.0332	0.041	0.0546	0.0766	0 0414	0.0397	0.0309	0.0447	0.0635
Cases of Agreement	112	115	106	119	105	107	104	90	120	
Cases of Disagreement	41	38	47	34	48	46	49	63	33	

the distribution of *individual* variations. To show the connection between individual variations and racial differences, I take all possible pairs of coefficients in the first nine columns of Table A and compare them with the corresponding pairs in the last column. When, in one of the first nine columns, the member of a pair which is larger is larger in the tenth column also, it counts one in the line at the bottom of Table A marked "Cases of Agreement." When, however, the larger member of a pair in one column is the smaller in the other, it counts one in the line marked "Cases of Disagreement." For example, in the column of Slavs, the *C. V.* of the head length is 0.0314, that of the head breadth is 0.0345, the head breadth being the more variable. In the column of racial differences also, the *C. V.* of the head breadth is greater than that of the head length; therefore this gives in column of Slavs one "Case of Agreement." On the other hand, for the Slavs, the heights of the forehead and of the nose do not agree in relative variability with the means of the twenty races given in the last column; consequently this comparison gives a "Case of Disagreement."

TABLE B.

Showing Order of Magnitude of C. V. of Table A.*

	24 Slavs.	8 Siamese.	12 Japanese.	20 Chinese.	9 Kanakas.	19 Jews.	20 Magyars.	26 Roumanians.	Average.	20 Races.
Head length	17	17	18	17	16	18	17	16	18	18
Head breadth	15	16	16	16	18	13	15	10	17	17
Nose length	5	8	6	3	13	8	1	2	3	3
Nose breadth	12	2	5	4	1	7	5	5	4	5
Nose height	2	1	2	1	9	3	3	3	2	1
Forehead height	1	5	1	2	2	1	2	1	1	2
Under jaw length	8	14	10	7	7	5	7	11	9	14
Mouth breadth	3	3	3	12	14	2	4	4	5	10
Upper face breadth	9	7	4	15	11	14	12	6	11	16
Lower face breadth	18	12	7	14	4	4	6	8.5	7	13
Upper arm length	7	9	9	9	10	9	11	13.5	10	6
Forearm length	11	15	17	11	17	16	9	7	16	15
Middle finger length	4	4	13	13	6	6	16	8.5	6	8
Hand width	10	11	15	10	8	11	14	18	14	4
Upper leg length	16	10	12	5	15	17	8	13.5	13	12
Lower leg length	14	18	8	6	5	12	13	12	12	11
Foot length	13	6	14	18	12	15	18	17	15	9
Foot breadth	6	13	11	8	3	10	10	15	8	7

* I. e. in column "20 Races," head length has C. V. smallest; head breadth, next larger; C. V. of nose height is largest.

It results, then, from the two lines of numbers at the bottom of Table A, that, in spite of the somewhat inaccurate values of this coefficient, in from two thirds to four fifths of the comparisons, that dimension in any pair which is the more variable among individuals of the same race is likewise the dimension which is the more diverse in the various races. Incidentally it may be noted, too, that the measurements of the face have, in general, the largest coefficients, and thus the high value for personal identification which has long been accorded to photographs of the face alone is justified.

TABLE C.

C. V. of Various Dimensions of Skulls of a Genus (Lepus) of Rodents. (Based on Coues and Allen, '77, pp. 222–226.)

	Mean dimensions of each of 12 Species.	15 Lepus campestris.	13 Lepus palustris.	Mean of the two Species.
Total length	0.121	0.0438	0.0178	0.0308
Greatest width	0.0935	0.0272	0.0239	0.0255
Distance between orbits	0.0825	0.0691	0.0475	0.0583
Nasal bones, length	0.122	0.0732	0.0348	0.0540
Nasal bones, width behind	0.150	0.0684	0.0327	0.0505
Nasal bones, width before	0.178	0.0760	0.0405	0.0582
Upper incisors to molars	0.125	0.0527	0.0280	0.0403
Upper incisors to hinder margin of palate	0.0799	0.0477	0.0238	0.0357
Upper incisors, height	0.146	0.0364	0.0568	0.0466
Upper incisors, width	0.146	0.0502	0.0434	0.0468
Length of upper molars	0.0958	0.0459	0.0393	0.0426
Width between upper molars	0.14	0.0348	0.0452	0.0400
Lower jaw length	0.124	0.0491	0.0240	0.0365
Lower jaw height	0.122	0.0329	0.0174	0.0251
Cases of Agreement		55	59	60
Cases of Disagreement		36	32	31

Table C shows for a genus of rodents what Table A shows for races of mankind. The first column is computed from the means of twelve species and subspecies of the genus Lepus. This gives, therefore, the relative value of these dimensions as specific and subspecific differences. The

remaining columns give the relative variability of the several dimensions in individuals of the two species. The numbers at the bottom of the table — "Cases of Agreement" and "Cases of Disagreement" — are found as in Table A. Table C is much less satisfactory than Table A, because the differences in variability between different dimensions are here small, and therefore the relative sizes of the coefficients are more at the mercy of accident. Even here, however, there is a decided preponderance of cases in which individual variation is directly correlated with specific difference.

TABLE D.

C. V. of Four Measurements of Two Genera of Rodents.
(Based on Miller, '93 and '93[a].)

	Tail.	Hind Foot.	Ear.	Tail and Body.
21 Zapus insignis	0.033	0.027	0.028	0.025
34 Zapus hudsonius	0.054	0.034	0.070	0.043
Mean	0.043	0.030	0.049	0.034
	Tail.	Hind Foot.	Ear.	Body.
105 Sitomys americanus	0.067	0.028	0.050	0.056
90 Sitomys americanus canadensis	0.070	0.031	0.052	0.049
Mean	0.068	0.029	0.051	0.053

Table D is computed from the measurements given with the description of two new species of rodents.* *Zapus insignis* and *Z. hudsonius* are reported to resemble each other closely, but to differ in the length of the ear. *Sitomys americanus* and *S. americanus canadensis* are described as nearly alike except in the length of the tail. Table D shows that, of the four dimensions which are recorded for these species, the mean *C. V.* of the ear is greatest in *Zapus*, and of the tail greatest in *Sitomys*. The *C. V.* in each single species, too, is greatest for the characteristic dimension, except in the case of *Z. insignis*. Possibly *Z. insignis* is the con-

* Miller ('93), pp. 1-8 and 55-70.

stant parent species from which *Z. hudsonius*, with its remarkably high coefficient for the ear, has separated. At any rate, in each genus, the characters which mark the species have the highest mean coefficient.

TABLE E.

C. V. of Measurements of Three Carnivora.

(Measured by the Author.)

	10 Lynx canadensis.	16 Felis domestica.	20 Vulpes fulvus.	Mean of Felis and Lynx.	Mean of 3 Species.
Length of nasal bones	0.050	0.054	0.039	0.052	0.048
Length of frontals	0.038	0.056	0.031	0.047	0.041
Length of snout, incisors to margin of palate .	0.039	0.062	0.033	0.050	0.045
Total basal length	0.028	0.054	0.029	0.041	0.037
Length from posterior nares to occipital foramen	0.024	0.058	0.026	0.041	0.036
Length of parietal and occipital bones . . .	0.030	0.046	0.027	0.038	0.034
Length of tooth series	0.044	0.050	0.024	0.047	0.039
Width between canines	0.041	0.055	0.045	0.048	0.047
Width between maxillary bones	0.045	0.040	0.028	0.042	0.038
Width of zygoma	0.049	0.054	0.044	0.051	0.049

Table E gives the *C. V.*'s of ten dimensions of the skulls of the cat, fox, and lynx. A comparison of the skulls of these three forms shows that the representatives of the cat and dog families are very similar, except for the length of the muzzle and the shape of the zygoma. A short muzzle and a wide zygoma are characteristic of all the cats,* and nowhere is the difference between the skulls more marked than in the length of the nasal bones. Examination of Table E shows that it is in just these dimensions — namely, those of the front part of the skull, the zygoma, and especially the nasal bones — that individuals are most variable; while, on the other hand, the "total basal length" and the dimensions of the back part of the skull, which are alike in the two families, give the smallest coefficients of variability.

* Flower ('70), p. 142.

Conclusion.

These four cases are the only ones to which I have applied this method of measuring variability. They are, however, taken entirely at random, and are in no wise selected cases. The conclusion to which they all point is that which, on general grounds, seems most likely to be true. This conclusion is, that there is so intimate a causal connection between the characters of individuals and those of the allied groups into which they are combined, that, in proportion as any character is variable in the individuals of one group, it is different in the allied groups.

Finally, I have to express my great indebtedness for assistance, criticism, and suggestion to Dr. Charles B. Davenport and Mr. Frederick H. Safford, both of Harvard University.

Wolfborough, N. H.,
March 5, 1896.

LITERATURE CITED.

Coues, E., and J. A. Allen.
'77. Monographs of North American Rodentia. Rept. U. S. Geol. Surv. Territories, Vol. XI. 1091 pp.

Flower, W. H.
'70. An Introduction to the Osteology of the Mammalia. 344 pp. London.

Galton, F.
'89. Natural Inheritance. 259 pp. London.

Merriman, M.
'84. A Text-book on the Method of Least Squares. New York.

Miller, G. S., Jr.
'93. A Jumping Mouse (Zapus insignis Miller) new to the United States. Proc. Biol. Soc. Washington, Vol. VIII. pp. 1–8. April 22.
'93[a]. Description of a new White-footed Mouse from the Eastern United States. Proc. Biol. Soc. Washington, Vol. VIII. pp. 55–70. June 20.

Quetelet, A.
'46. Lettres sur la Théorie des Probabilités, appliquée aux sciences morales et politiques. 450 pp. Bruxelles.

Stieda, L.
'82. Ueber die Anwendung der Wahrscheinlichkeitsrechnung in der anthropologischen Statistik. Arch. f. Anthrop., Band XIV. pp. 167–182.

Weisbach, A.
'78. Körpermessungen verschiedener Menschenrassen. Berlin.

Weldon, W. F. R.
'92. Certain Correlated Variations in *Crangon vulgaris*. Proc. Roy. Soc. London, Vol. LI. pp. 2–21.